Probability Comprehensive Exam
Spring 2019

Student Number: ☐

Instructions: Complete 5 of the 9 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8 9

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Let X be a non-negative random variable, such that $0 < \mathbb{E}X < +\infty$, and let $0 < x < 1$. Show that
$$\mathbb{P}(X \geq x\mathbb{E}X) \geq (1-x)^2 \frac{(\mathbb{E}X)^2}{\mathbb{E}(X^2)}.$$

Solution:

If $\mathbb{E}(X^2) = +\infty$, then the inequality is trvially true. So without loss of generality, assume that $\mathbb{E}X^2 < +\infty$, in which case necessarily $\mathbb{E}X^2 > 0$, since $\mathbb{E}X > 0$. Next,
$$\mathbb{E}X = \mathbb{E}X\mathbf{1}_{X \geq x\mathbb{E}X} + \mathbb{E}X\mathbf{1}_{X < x\mathbb{E}X} \leq \mathbb{E}X\mathbf{1}_{X \geq x\mathbb{E}X} + x\mathbb{E}X.$$
Therefore,
$$(1-x)\mathbb{E}X \leq \mathbb{E}X\mathbf{1}_{X \geq x\mathbb{E}X} \leq \sqrt{\mathbb{P}(X \geq s\mathbb{E}X)}\sqrt{\mathbb{E}(X^2)},$$
by the Cauchy-Schwarz inequality. This proves the result.

2. If $(X_n)_{n\geq 1}$ is a sequence of random variables, then there exists a sequence $(c_n)_{n\geq 1}$ with $c_n \to \infty$, such that
$$\mathbb{P}(\lim_{n\to\infty} \frac{X_n}{c_n} = 0) = 1.$$

Solution: By Borel-Cantelli, it suffices to choose c_n such that
$$\sum_{n\geq 1} \mathbb{P}(|X_n| > \epsilon c_n) < \infty$$
for all choices of $\epsilon > 0$. We can choose first d_n such that $\mathbb{P}(|X_n| > d_n) < 1/2^n$ for each n. This is possible because for each n, $\mathbb{P}(|X_n| > \lambda) \xrightarrow[\lambda\to\infty]{} 0$, thus we can choose for each n a $\lambda_n > 0$ such that
$$\mathbb{P}(|X_n| > \lambda_n) < 1/2^n.$$
Using this, take $d_n = \lambda_n$. However to choose c_n we will take them such that $c_n = \max\{n, \max_{k=1,2,\dots,n} d_k\}$. Clearly now, c_n is increasing to infinity and $c_n \geq d_n$. It remains to observe that for small $\epsilon > 0$
$$\sum_{n\geq 1} \mathbb{P}(|X_n| \geq \epsilon c_n) \leq \sum_{n\geq 1} \mathbb{P}(|X_n| \geq c_n) < \infty$$
and from this we get $|X_n|/c_n$ converges to 0 a.s.

3. Assume that $\{X_n\}_{n\geq 1}$ are random variables such that

   1. $E[X_n] = 0$ and $\mathbb{E}[X_n^2] \leq 1$ for any $n \geq 1$
   2. $\mathbb{E}[X_i X_j] \leq 0$ for any $i \neq j$.

   Show that for any sequence $\{a_n\}_{n\geq 1} \subset [1/2, 2]$,
$$\frac{a_1 X_1 + a_2 X_2 + \cdots + a_n X_n}{a_1 + a_2 + \cdots + a_n} \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Solution: We use the standard proof of the weak law of large numbers for the case of finite variance. Denote $\bar{S}_n = \frac{a_1X_1+a_2X_2+\cdots+a_nX_n}{a_1+a_2+\cdots+a_n}$ and use Chebyshev's inequality to justify that
$$\mathbb{P}(|\bar{S}_n| \geq \epsilon) \leq \frac{\mathbb{E}[\bar{S}_n^2]}{\epsilon^2}.$$
Now,
$$\mathbb{E}[(\sum_{i=1}^n a_i X_i)^2] = \sum_{i,j=1}^n a_i^2 a_j^2 \mathbb{E}[X_i X_j] \leq \sum_{i=1}^n a_i^2 \mathbb{E}[X_i^2] \leq \sum_{i=1}^n a_i^2.$$
Thus we get that
$$\frac{\mathbb{E}[\bar{S}_n^2]}{\epsilon^2} \leq \frac{\sum_{i=1}^n a_i^2}{\epsilon^2 (\sum_{i=1}^n a_i)^2} \leq \frac{4n}{\epsilon^2 (n/2)^2} = \frac{16}{n\epsilon^2}.$$
Which proves the claim.

4. Let $(X_n)_{n\geq 1}$ be a sequence of non -negative uniformly integrable random variables such that, as $n \to +\infty$, $X_n \Rightarrow X$. Show that X is integrable and that $\lim_{n\to+\infty} \mathbb{E}X_n = \mathbb{E}X$.

Solution: By weak convergence, $\liminf_{n\to+\infty} \mathbb{P}(X_n > t) = \mathbb{P}(X > t)$, except possibly at countably many t, while by uniform integrability, the sequence $(\mathbb{E}X_n)_{n\geq 1}$ is bounded. Hence, by Fatou's Lemma,
$$\begin{aligned}\mathbb{E}X = \int_0^{+\infty} \mathbb{P}(X > t)dt &= \int_0^{+\infty} \liminf_n \mathbb{P}(X_n > t)dt \\ &\leq \liminf_n \int_0^{+\infty} \mathbb{P}(X_n > t)dt \\ &= \liminf_n \mathbb{E}X_n < +\infty.\end{aligned}$$
Next, for any $M > 0$,
$$\mathbb{E}X_n = \int_0^M \mathbb{P}(X_n > t)dt + \mathbb{E}X_n \mathbf{1}_{X_n \geq M} = \int_0^M \mathbb{P}(M > X_n > t)dt + \mathbb{E}X_n \mathbf{1}_{X_n \geq M},$$

and similarly,

$$\mathbb{E}X = \int_0^M \mathbb{P}(M > X > t)dt + \mathbb{E}X_n \mathbf{1}_{X \geq M}.$$

By uniform integrability, for each $\epsilon > 0$, there is an $M > 0$ such that the last terms in each one of the above equalities are less than ϵ. So, to conclude it is enough to show that as $n \to +\infty$, $\int_0^M \mathbb{P}(M > X_n > t)dt$ converges to $\int_0^M \mathbb{P}(M > X > t)dt$. But, since M can be chosen in such a way that $\mathbb{P}(X = M) = 0$, weak convergence and dominated convergence on $[0, M]$ give the conclusion.

5. If $X_1, X_2, \ldots, X_n$ are iid exponential random variables with parameter 1, compute the almost sure limit of

$$\frac{1}{n}\sum_{i=1}^{n} e^{-X_k - 2X_{k+1} - 3X_{k+2}}$$

as n tends to infinity.

Solution: We split this according to

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} e^{-X_k-2X_{k+1}-3X_{k+2}} =& \frac{1}{n}\sum_{i=0}^{[n-3)/3]} e^{-X_{3i+1}-2X_{3i+2}-3X_{3i+3}} \\ &+ \frac{1}{n}\sum_{i=0}^{[(n-4)/3]} e^{-X_{3i+2}-2X_{3i+3}-3X_{3i+4}} \\ &+ \frac{1}{n}\sum_{i=0}^{[(n-5)/3]} e^{-X_{3i+3}-2X_{3i+4}-3X_{3i+5}} \\ &+ \frac{R_n}{n} \end{aligned}$$

where R_n is eventually a remainder which is certainly less than 2. Now using the strong law of large numbers, for each sum $\frac{1}{n}\sum_{i=0}^{[n-3)/3]} e^{-X_{3i+1}-2X_{3i+2}-3X_{3i+3}}$, $\frac{1}{n}\sum_{i=0}^{[(n-4)/3]} e^{-X_{3i+2}-2X_{3i+3}-3X_{3i+4}}$, $\frac{1}{n}\sum_{i=0}^{[(n-5)/3]} e^{-X_{3i+3}-2X_{3i+4}-3X_{3i+5}}$, we get that in almost sure sense, the limit is

$$\mathbb{E}[e^{X_1-2X_2-3X_3}] = \mathbb{E}[e^{-X_1}]\mathbb{E}[e^{-2X_2}]\mathbb{E}[e^{-3X_3}] = \int_0^1 e^{-2x}dx \int_0^1 e^{-3x}dx \int_0^1 e^{-4x}dx = \frac{1}{24}.$$

6. Assume $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space such that there exist $X_1, X_2 : \Omega \to \mathbb{R}$ two independent Bernoulli random variables such that $\mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = 1/2$. Show that Ω must have at least 4 elements.

   Give an example with Ω having 4 elements together with a sigma algebra such that on it we can define two independent Bernoulli as above.

   Can you generalize this?

   Solution: Since $\mathcal{F}$ has at least 4 disjoint events, namely, $\{X_1 = 0, X_2 = 0\}$, $\{X_1 = 0, X_2 = 1\}$, $\{X_1 = 1, X_2 = 0\}$, $\{X_1 = 1, X_2 = 1\}$, it follows that in fact $\mathcal{F}$ contains more than 2^4 elements, for example, all sets which are disjoint unions of these four elements (also including the empty union) gives at least $2^4 = 16$ elements. Since $\mathcal{F} \subset \mathcal{P}(\Omega)$, it means that Ω must have at least 4 elements, otherwise $\mathcal{P}(\Omega)$ has at most 2^3 elements.

   On $\Omega = \{0,1\} \times \{0,1\}$ and the sigma algebra of all subsets, we can define $X_1(\omega_1, \omega_2) = \omega_1$ and $X_2(\omega_1, \omega_2) = \omega_2$. This is the standard tensor product construction.

   For a generalization, if we have n independent Bernoulli random variables then Ω must have at least 2^n elements. Indeed, we have 2^n disjoint subsets in $\mathcal{F}$ and thus $\mathcal{F}$ must have 2^{2^n} elements. This implies that Ω must have at least 2^n elements.

7. If X, Y are two random variables such that $X \geq Y$ and X, Y have the same distribution, then $X = Y$ almost surely.

   Solution: We try to relate the cumulative functions. Thus

$$\mathbb{P}(X \leq x) = \mathbb{P}(Y \leq X \leq x) \leq \mathbb{P}(Y \leq x, X \leq x) = \mathbb{P}(Y \leq x) - \mathbb{P}(Y \leq x < X)$$

   Thus, because $\mathbb{P}(Y \leq x) = \mathbb{P}(X \leq x)$, we get that $\mathbb{P}(Y \leq x < X) = 0$ for any choice of $x \in \mathbb{R}$. Finally,

$$\mathbb{P}(Y < X) \leq \sum_{r \text{ rational}} \mathbb{P}(Y \leq r < X) = 0.$$

   Consequently, $\mathbb{P}(Y = X) = 1$.

8. Assume that $X_1, X_2, \ldots, X_n$ are iid with density $f(x) = \frac{2}{x^3}$ for $x \geq 1$ and 0 otherwise. Define

$$M_n = \frac{1}{n} \max\{X_1, \sqrt{2}X_2, \ldots, \sqrt{n}X_n\}.$$

   Show that X_n converges in distribution and find the limit.

Solution: We compute the cumulative function as for $x > 0$

$$F_{M_n}(x) = \mathbb{P}(X_1 \le nx, X_2 \le nx/\sqrt{2}, \ldots, X_n \le nx/\sqrt{n}) = \prod_{k=1}^{n} F_X(n^2x/k).$$

Now the cumulative function of X is (for $x \ge 1$)

$$F_X(x) = \int_1^x \frac{2}{t^3} dt = 1 - 1/x^2.$$

Thus we have for $x > 0$ and large n, that

$$F_{M_n}(x) = \prod_{k=1}^{n} (1 - \frac{k}{n^2x^2})$$

To compute the limit of this we take the log and use the fact that

$$\ln(1-t) = -t + O(t^2)$$

for small, t, thus

$$\ln(F_{M_n}) = \sum_{k=1}^{n} \ln(1 - \frac{k}{n^2x^2}) \approx -\frac{\sum_{k=1}^{n} k}{n^2x^2} \xrightarrow[n\to\infty]{} -\frac{1}{2x^2}.$$

Therefore,

$$F_{M_n}(x) \xrightarrow[n\to\infty]{} F(x) = e^{-1/(2x^2)}, \text{ for } x > 0.$$

9. Let X be a finite mean random variable, let $\mathbf{F}$ be a σ-field and let G be a σ-field independent of $\sigma(\sigma(X), \mathbf{F})$. (As usual, $\sigma(X)$ is the σ-field generated by X and $\sigma(\sigma(X), \mathbf{F})$ is the σ-field generated by $\sigma(X)$ and $\mathbf{F}$.) Is it true or false that $\mathbb{E}(X|\sigma(\mathbf{F}, \mathbf{G})) = \mathbb{E}(X|\mathbf{F})$?

Solution: Yes it is true! Let $F \in \mathbf{F}$ and let $G \in \mathbf{G}$, then $F \cap G \in \sigma(\mathbf{F}, \mathbf{G})$ and using the very definition of conditional expectation as well as independence (twice) we get:

$$\begin{aligned}\mathbb{E}[\mathbb{E}(X|\sigma(\mathbf{F},\mathbf{G}))\mathbf{1}_{F\cap G}] &= \mathbb{E}(X\mathbf{1}_{F\cap G}) = \mathbb{E}(X\mathbf{1}_F\mathbf{1}_G) = \mathbb{E}(X\mathbf{1}_F)\mathbb{E}\mathbf{1}_G \\ &= \mathbb{E}[\mathbb{E}(X|\mathbf{F})\mathbf{1}_F)]\mathbb{E}\mathbf{1}_G = \mathbb{E}[\mathbb{E}(X|\mathbf{F})\mathbf{1}_F)\mathbf{1}_G] = \mathbb{E}[\mathbb{E}(X|\mathbf{F})\mathbf{1}_{F\cap G}].\end{aligned}$$

Therefore $\mathbb{E}(X|\sigma(\mathbf{F},\mathbf{G}))$ and $\mathbb{E}(X|\mathbf{F})$ agree on a π-system generating $\sigma(\mathbf{F},\mathbf{G})$. Now, let μ_1 and μ_2 be respectively defined via $\mu_1(A) = \mathbb{E}[\mathbb{E}(X|\sigma(\mathbf{F},\mathbf{G}))\mathbf{1}_A]$ and $\mu_2(A) =$

$\mathbb{E}[\mathbb{E}(X|\mathbf{F})\mathbf{1}_A]$. Since $\mathbb{E}|X| < +\infty$, then μ_1 and μ_2 are finite measures which agree on a π-system generating $\sigma(\mathbf{F}, \mathbf{G})$, so they must agree on $\sigma(\mathbf{F}, \mathbf{G})$. Finally, by uniqueness,

$$\mathbb{E}(X|\sigma(\mathbf{F}, \mathbf{G})) = \mathbb{E}(X|\mathbf{F}).$$

This proves the result. (Above, instead of μ_1 and μ_2, one could also consider positive measures by looking at the positive and the negative part of X).